

PRELIMINARY AMENDMENT

Serial Number:

Filing Date:

Title: SYSTEM AND METHOD FOR CONTENT-BASED PARTITIONING AND MINING

Assignee: Intel Corporation

Page 2

Dkt: P23151 (INTEL)

In the claims

Please add claims 21-23 as detailed below:

1. (Original) A system for mining data comprising:
a data store including data having a number of items;
a mining application to mine data in the data store, the mining application including logic, the logic, when executed, is to:
identify a number of frequent items of the data store;
compute a probe structure based on the number of identified frequent items; and,
partition the data according to content of the probe structure;
wherein the mining application uses the probe structure to build a frequent pattern tree (FP-tree); and
a memory for storing the probe structure and the FP-tree.
2. (Original) The system of claim 1, wherein the data of the data store includes a number of transactions, wherein each transaction comprises a unique sequence of items identified by the logic when identifying the frequent items of the data store.
3. (Original) The system of claim 2, wherein the logic is to partition the transactions according to content of the identified frequent items to obtain the probe structure, wherein the probe structure includes combinations of the identified frequent items and the number of occurrences of one or more content-based transactions.
4. (Original) The system of claim 3, wherein the logic orders the identified frequent items based on an occurrence frequency of each identified item in the data store.

-
5. (Original) The system of claim 3, further comprising a heuristic algorithm, wherein the heuristic algorithm is to group the one or more content-based transactions into approximately equal groups.
6. (Original) The system of claim 1, further comprising a master processor and one or more slave processors, wherein the master processor is to distribute a group of transactions to the one or more slave processors to build the FP-tree.
7. (Original) The system of claim 6, wherein the one or more slave processors build a part of the FP-tree based on the grouping of content-based transactions.
8. (Original) The system of claim 7, wherein the multiple processors mine the FP-tree to determine unique information about the items of the data store.
9. (Original) The system of claim 1, further comprising a multi-core system architecture.
10. (Original) A system for mining data, the system comprising:
a database including a number of transactions;
at least one processor to perform mining operations on the database, the at least one processor is to execute content-based partitioning logic on the transactions, wherein the content-based partitioning logic is to partition the transactions according to content based on a number of identified frequent items to obtain a probe structure; and
a memory to store the probe structure.
11. (Original) The system of claim 10, the probe structure further comprising a probe tree and probe table, wherein the probe tree and probe table further comprise 2^M branches, wherein M corresponds to the number of identified frequent items.

12. (Original) The system of claim 11, wherein the memory further comprises shared memory to store the probe tree and probe table.

13. (Original) The system of claim 11 further comprising multiple processors to recursively mine the database, wherein each processor shares a substantially equal load based on a grouping and distribution of the 2^M branches.

14. (Original) The system of claim 13, the multiple processors further comprising a master processor and at least one slave processor to perform mining operations, wherein the master processor distributes operations to the at least one slave processor when building a frequent pattern tree (FP-tree) using the probe structure.

15. (Original) A method for mining data of a database, comprising:
identifying frequent items of the database;
building a probe structure based on the identified frequent items, wherein each branch of the probe structure includes a number of identified frequent items based on content;
grouping the branches of the probe structure based on the content of each branch;
and
building a frequent pattern tree (FP-tree) from the probe structure.

16. (Original) The method of claim 15, further comprising scanning a first portion of the database when identifying frequent items of the database, and scanning a second portion of the database when building the probe structure, wherein the probe structure includes an associated number of counts with each branch of the probe structure after scanning the second portion of the database.

17. (Original) The method of claim 15, further comprising building the probe structure to include a probe tree and probe table, and using the probe tree and probe table to build the FP-tree for mining the FP-tree to determine frequent data patterns.

18. (Original) The method of claim 15, further comprising distributing each group of branches to an associated processor before building the FP-tree.

19. (Original) The method of claim 18, further comprising using a master processor to distribute each group of branches to one or more slave processors, and using the one or more slave processors to build the FP-tree.

20. (Original) The method of claim 15, further comprising partitioning the database according to content of the identified frequent items to obtain the probe structure, wherein the probe structure includes combinations of the identified frequent items and the number of occurrences of one or more content-based transactions.

21. (New) A computer-readable medium having stored thereon instructions, which when executed in a system operate to manage data of a database by:

identifying frequent items of the database;

building a probe structure based on the identified frequent items, wherein each branch of the probe structure includes a number of identified frequent items assorted by content;

grouping the branches of the probe structure based on the content of each branch;

and

building a frequent pattern tree (FP-tree) from the probe structure.

22. (New) The computer-readable medium of claim 21, wherein the instructions, which when executed in a system operate to manage data of a database further by

building the probe structure to include a probe tree and probe table, and using the probe tree and probe table to build the FP-tree for mining the FP-tree to determine frequent data patterns.

23. (New) The computer-readable medium of claim 21, wherein the instructions, which when executed in a system operate to manage data of a database further by distributing each group of branches to an associated processor before building the FP-tree.